ELSEVIER

Journal of Econometrics 110 (2002) 417–435

JOURNAL OF Econometrics

www.elsevier.com/locate/econbase

Evaluating GARCH models

Stefan Lundbergh, Timo Teräsvirta*

Department of Economic Statistics, Stockholm School of Economics, P.O. Box 6501, SE-113 83 Stockholm, Sweden

Abstract

In this paper, a unified framework for testing the adequacy of an estimated GARCH model is presented. Parametric Lagrange multiplier (LM) or LM type tests of no ARCH in standardized errors, linearity, and parameter constancy are proposed. The asymptotic null distributions of the tests are standard, which makes application easy. Versions of the tests that are robust against nonnormal errors are provided. The finite sample properties of the test statistics are investigated by simulation. The robust tests prove superior to the nonrobust ones when the errors are nonnormal. They also compare favourably in terms of power with misspecification tests previously proposed in the literature.

JEL classification: C22; C52

Keywords: Conditional heteroskedasticity; Model misspecification test; Nonlinear time series; Parameter constancy; Smooth transition GARCH

1. Introduction

When modelling the conditional mean, at least when it is a linear function of parameters, the estimated model is regularly subjected to a battery of misspecification tests to check its adequacy. The hypothesis of no (conditional) heteroskedasticity, no error autocorrelation, linearity, and parameter constancy, to name a few, are tested using various methods. As for models of conditional variance, such as the popular GARCH model, testing the adequacy of the estimated model has been much less common in practice. But then, misspecification tests do exist for GARCH models as well. For example, Bollerslev (1986) already suggested a score or Lagrange multiplier (LM) test for testing a GARCH model against a higher order GARCH model. Li and Mak (1994)

* Corresponding author.
E-mail address: sttimo@hhs.se (T. Teräsvirta).

PII: S0304-4076(02)00096-9

derived a portmanteau type test for testing the adequacy of a GARCH model. Engle and Ng (1993) considered testing the GARCH specification against asymmetry using the so-called sign-bias and size-bias tests. Chu (1995) derived a test of parameter constancy against a single structural break. This test has a nonstandard asymptotic null distribution, but Chu provided tables for critical values. Recently, Lin and Yang (1999) derived another test against a single structural break, based on empirical distribution functions.

In this article, we present a number of simple misspecification tests for GARCH models. The idea is to make misspecification testing easy without sacrificing power. We consider testing the null of no ARCH in the standardized errors, linearity or symmetry against a smooth transition GARCH, and parameter constancy against smoothly changing parameters. A single structural break is nested in the alternative hypothesis of our parameter constancy test. A two-regime asymmetric GARCH model such as the so-called GJR model (Glosten et al., 1993) is nested in the alternative of smooth transition GARCH. Note that the LM test of Bollerslev (1986) fits into our framework. Furthermore, we show that under normality the portmanteau test of Li and Mak (1994) is asymptotically equivalent to our test of no remaining ARCH. All our tests are LM or LM type tests and require only standard asymptotic distribution theory. They may be obtained from the same "root" by merely changing the definitions of the elements of the score vector corresponding to the alternative hypothesis.

Very often in applications, the assumption of a normal error distribution of a GARCH process is too restrictive. A useful feature of our tests is that robustifying them against nonnormal errors is straightforward. In fact, our Monte Carlo experiments suggest that one should always apply the robust versions of our tests. At the sample size 1000, used in our simulation study and rather typical for GARCH applications, the efficiency loss compared to the nonrobust tests appears to be minimal when the errors are normal. On the other hand, the same simulations show that the nonrobust tests tend to be undersized if the error distribution is leptokurtic and that, in this situation, robust tests have clearly better power than the nonrobust ones. In terms of power, the tests proposed in this paper compare favourably with corresponding tests currently available in the literature. This is true for all three cases we consider: no remaining ARCH, linearity, and parameter constancy.

The plan of the paper is as follows. In Section 2, we define the model. In Section 3, we discuss testing the null of no ARCH in the standardized errors and compare our LM-test with the portmanteau test of Li and Mak (1994). Section 4 considers testing null hypotheses of linearity and parameter constancy. Section 5 contains results of a simulation experiment and Section 6 concludes.

2. The model

Consider a time series model with the following structure:

$$y_t = f(\mathbf{w}_t; \boldsymbol{\varphi}) + \varepsilon_t, \tag{2.1}$$

where f is at least twice continuously differentiable with respect to $\boldsymbol{\varphi} \in \boldsymbol{\Phi}$, for all $\mathbf{w}_t = (\mathbf{y}_{t-1}', \mathbf{u}_t')'$ with $\mathbf{y}_{t-1} = (1, y_{t-1}, \ldots, y_{t-n})' \in \mathfrak{R}^{n+1}$ and exogenous $\mathbf{u}_t = (u_{1t}, \ldots, u_{kt})' \in \mathfrak{R}^k$, everywhere in $\mathbf{\Phi}$. The error is parameterized as

$$\varepsilon_t = \xi_t h_t^{1/2}, \tag{2.2}$$

where $\{\xi_t\}$ is a sequence of independent identically distributed random variables with mean zero, unit variance and $\mathsf{E}\xi_t^3 = 0$. Furthermore, the conditional variance $h_t = \mathbf{\eta}'\mathbf{s}_t$ such that $\mathbf{s}_t = (1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-q}^2, h_{t-1}, \ldots, h_{t-p})'$ and $\mathbf{\eta} = (\alpha_0, \alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p)'$ with $\alpha_0 > 0$, whereas $\alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p$ satisfy the conditions in Nelson and Cao (1992) that ensure the positivity of h_t. These conditions allow some of the parameters to be negative, unless $p = q = 1$. Model (2.2) is thus the standard GARCH(p,q) model. We assume regularity conditions hold such that the central limit theorem and the law of large numbers apply whenever required. For such conditions in the multivariate GARCH(p,q) case, see Comte and Lieberman (2000). In the univariate case, their conditions require the density of ξ_t to be absolutely continuous with respect to the Lebesgue measure and positive in a neighbourhood of the origin. Furthermore, it is required that $\mathsf{E}\varepsilon_t^8 < \infty$, which of course implies further restrictions on the density of ξ_t.

The assumption $\mathsf{E}\xi_t^3 = 0$ that Comte and Lieberman (2000) do not need guarantees block diagonality of the information matrix of the log-likelihood function. However, it is not just a technical simplification. We shall consider, among other things, a test that has power against asymmetric response to shocks. In deriving such a test it is appropriate to assume that the conditional distribution of ε_t given h_t, is not skewed.

3. Testing the adequacy of the GARCH model

3.1. Testing the null hypothesis of no remaining ARCH

In order to consider the adequacy of the GARCH model, we formulate a parametric alternative to the model. Assume that in (2.2),

$$\xi_t = z_t g_t^{1/2}, \tag{3.1}$$

where $\{z_t\}$ is a sequence of independent, identically distributed random variables with zero mean, unit variance and $\mathsf{E}z_t^3 = 0$. Furthermore, $g_t = 1 + \mathbf{\pi}'\mathbf{v}_t$, where $\mathbf{v}_t = (\xi_{t-1}^2, \ldots, \xi_{t-m}^2)'$ and $\mathbf{\pi} = (\pi_1, \ldots, \pi_m)'$, $\pi_j \geqslant 0$, $j = 1, \ldots, m$. Eq. (3.1) may thus be written as

$$\varepsilon_t = z_t (h_t g_t)^{1/2} \tag{3.2}$$

and could be called an "ARCH nested in GARCH" model as $\xi_{t-j}^2 = \varepsilon_{t-j}^2 / h_{t-j}$, $j = 1, \ldots, m$. We want to test $\mathrm{H}_0\colon \mathbf{\pi} = \mathbf{0}$ against $\mathbf{\pi} \neq \mathbf{0}$ and thus follow the standard practice of choosing a two-sided alternative although the elements of $\mathbf{\pi}$ are constrained to be non-negative. Under this hypothesis, $g_t \equiv 1$, and the model collapses into a GARCH(p,q) model. (For ways of testing $\mathrm{H}_0\colon \mathbf{\pi} = \mathbf{0}$ against $\mathbf{\pi} > \mathbf{0}$ when $h_t \equiv \alpha_0$, see Lee and King (1993) and Demos and Sentana (1998).)

We introduce the following notation. Let $\hat{\varepsilon}_t$ and $\hat{h}_t$ be the error ε_t and the conditional variance h_t, respectively, estimated under H_0, and $\hat{\xi}_t^2 = \hat{\varepsilon}_t^2 / \hat{h}_t$. Furthermore, let $\hat{\mathbf{x}}_t = \hat{h}_t^{-1}\partial\hat{h}_t/\partial\boldsymbol{\eta}'$ ($\partial\hat{h}_t/\partial\boldsymbol{\eta}'$ denotes $\partial h_t/\partial\boldsymbol{\eta}'$ estimated under H_0) and $\hat{\mathbf{v}}_t = (\hat{\xi}^2_{t-1},\ldots,\hat{\xi}^2_{t-m})'$. The quasi maximum likelihood approach leads to the following result:

Theorem 3.1. *Consider the model* (3.2) *where* $g_t = 1 + \boldsymbol{\pi}'\mathbf{v}_t$ *and* $\{z_t\}$ *is a sequence of independent identically distributed random variables with zero mean, unit variance and* $\mathsf{E}z_t^3 = 0$. *Under* $\mathrm{H}_0\colon \boldsymbol{\pi} = \mathbf{0}$, the statistic

$$LM_\pi = (1/4T)\left\{\sum_{t=1}^{T}(\hat{\varepsilon}_t^2/\hat{h}_t - 1)\hat{\mathbf{v}}_t'\right\}\mathbf{V}(\hat{\boldsymbol{\eta}})^{-1}\left\{\sum_{t=1}^{T}(\hat{\varepsilon}_t^2/\hat{h}_t - 1)\hat{\mathbf{v}}_t\right\}, \tag{3.3}$$

where

$$V(\hat{\boldsymbol{\eta}})^{-1} = (4T/\hat{k})\left\{\sum_{t=1}^{T}\hat{\mathbf{v}}_t\hat{\mathbf{v}}_t' - \sum_{t=1}^{T}\hat{\mathbf{v}}_t\hat{\mathbf{x}}_t'\left(\sum_{t=1}^{T}\hat{\mathbf{x}}_t\hat{\mathbf{x}}_t'\right)^{-1}\sum_{t=1}^{T}\hat{\mathbf{x}}_t\hat{\mathbf{v}}_t'\right\}^{-1}$$

with $\hat{k}=(1/T)\sum_{t=1}^{T}(\hat{\varepsilon}_t^2/\hat{h}_t-1)^2$ *is a consistent estimator of the inverse of the covariance matrix of the partial score under the null hypothesis, has an asymptotic* χ^2 *distribution with* m *degrees of freedom.*

Proof. See the appendix. □

The test may also be carried out using an artificial regression as follows.

1. Estimate the parameters of the conditional variance model under the null, compute $\hat{\varepsilon}_t^2/\hat{h}_t - 1, t = 1,\ldots,T$, and $SSR_0 = \sum_{t=1}^{T}(\hat{\varepsilon}_t^2/\hat{h}_t - 1)^2$.
2. Regress $(\hat{\varepsilon}_t^2/\hat{h}_t - 1)$ on $\hat{\mathbf{x}}_t'$, $\hat{\mathbf{v}}_t'$ and compute the sum of squared residuals, SSR_1.
3. Compute the test statistic $LM = T(SSR_0 - SSR_1)/SSR_0$ or the F-version:

$$F = \frac{(SSR_0 - SSR_1)/m}{SSR_1/(T - p - q - 1 - m)}.$$

For the sample sizes relevant in GARCH modelling, there is no essential difference between the properties of LM and F.

If the standard GARCH parameterization is assumed to be the true data-generating process, then it is well-known in practice that the iid error process often cannot be assumed normal, although the likelihood is constructed under the assumption of normality. It is therefore desirable to make the test robust against nonnormal errors. The robust version of the LM test can be constructed following Wooldridge (1991, Procedure 4.1). The test is carried out as follows:

1. Obtain the quasi maximum likelihood estimate for $\boldsymbol{\eta}$ under H_0, compute $\hat{\varepsilon}_t^2/\hat{h}_t - 1, \hat{\mathbf{v}}_t$ and $\hat{\mathbf{x}}_t,\ t = 1,\ldots,T$.
2. Regress $\hat{\mathbf{v}}_t$ on $\hat{\mathbf{x}}_t$, and compute the $(m \times 1)$ residual vectors $\mathbf{r}_t, t = 1,\ldots,T$.

3. Regress 1 on $(\hat{\varepsilon}_t^2/\hat{h}_t - 1)\mathbf{r}_t$ and compute the residual sum of squares SSR from that regression. The test statistic

$$LM_\pi^{\mathrm{R}} = T - SSR$$

has an asymptotic χ^2 distribution with m degrees-of-freedom under the null hypothesis.

The purpose of the second step is to purge the effects of the normality assumption from $\hat{\mathbf{v}}_t$. The resulting test statistic has the same asymptotic distribution as (3.3) and is, under normality, asymptotically efficient; see Wooldridge (1991). Any other consistent estimator of $\boldsymbol{\eta}$ may be employed in place of the quasi-maximum likelihood one.

3.2. A portmanteau test and a comparison

Li and Mak (1994) recently introduced a portmanteau statistic for testing the adequacy of the standard GARCH(p,q) model. The null hypothesis is that the squared and standardized error process is not autocorrelated. Let $\mathbf{r}=(r_1,\dots,r_m)'$ be the $m\times 1$ vector of the first m autocorrelations of $\{\varepsilon_t^2/h_t\}$ so that H_0:$\mathbf{r}=\mathbf{0}$. Assuming normal errors, Li and Mak (1994) showed that under this hypothesis $\sqrt{T}\hat{\mathbf{r}}$ is asymptotically normally distributed where T is the number of observations and $\hat{\mathbf{r}}=(1/2T)\sum(\hat{\varepsilon}_t^2/\hat{h}_t-1)\hat{\mathbf{v}}_t^*$ with $\hat{\varepsilon}_t=y_t-f(\mathbf{w}_t;\hat{\boldsymbol{\varphi}})$, $\hat{h}_t=h_t(\mathbf{w}_t;\hat{\boldsymbol{\varphi}},\hat{\boldsymbol{\eta}})$ and $\hat{\mathbf{v}}_t^*=(\hat{\varepsilon}_{t-1}^2/\hat{h}_{t-1}-1,\dots,\hat{\varepsilon}_{t-m}^2/\hat{h}_{t-m}-1)'$. Furthermore, under the null hypothesis the asymptotic covariance matrix of $\sqrt{T}\hat{\mathbf{r}}$ has the form

$$\mathbf{V}_{\mathbf{r}}(\hat{\boldsymbol{\eta}})=\mathbf{I}_m-\mathbf{X}_{\mathbf{r}}(\hat{\boldsymbol{\eta}})'\mathbf{G}^{-1}(\hat{\boldsymbol{\eta}})\mathbf{X}_{\mathbf{r}}(\hat{\boldsymbol{\eta}}) \tag{3.4}$$

as $\mathrm{plim}_{T\to\infty}T^{-1}\sum\hat{\mathbf{v}}_t^*\hat{\mathbf{v}}_t^{*'}=2\mathbf{I}_m$ under H_0, and $\mathbf{X}_{\mathbf{r}}(\hat{\boldsymbol{\eta}})=-(1/2T)\sum\hat{\mathbf{x}}_t\hat{\mathbf{v}}_t^{*'}$. In (3.4), $\mathbf{G}^{-1}(\hat{\boldsymbol{\eta}})$ is a consistent estimator of the relevant block of the information matrix, evaluated at $\boldsymbol{\eta}=\hat{\boldsymbol{\eta}}$. The portmanteau statistic becomes

$$Q(m)=T\hat{\mathbf{r}}'\mathbf{V}_{\mathbf{r}}(\hat{\boldsymbol{\eta}})^{-1}\hat{\mathbf{r}} \tag{3.5}$$

which, as Li and Mak (1994) showed, is asymptotically χ^2-distributed with m degrees of freedom under the null hypothesis. In order to link (3.5) to our theory in the case of normal errors, define

$$Q^*(m)=\frac{1}{4T}\left(\sum(\hat{\varepsilon}_t^2/\hat{h}_t-1)\hat{\mathbf{v}}_t^{*'}\right)\mathbf{V}_{\mathbf{r}}^*(\hat{\boldsymbol{\eta}})^{-1}\left(\sum\hat{\mathbf{v}}_t^*(\hat{\varepsilon}_t^2/\hat{h}_t-1)\right), \tag{3.6}$$

where $\mathbf{V}_{\mathbf{r}}^*(\hat{\boldsymbol{\eta}})=(1/2T)(\sum\hat{\mathbf{v}}_t^*\hat{\mathbf{v}}_t^{*'}-\sum\hat{\mathbf{v}}_t^*\hat{\mathbf{x}}_t'(\sum\hat{\mathbf{x}}_t\hat{\mathbf{x}}_t')^{-1}\sum\hat{\mathbf{x}}_t\hat{\mathbf{v}}_t^{*'})$. The only difference between (3.5) and (3.6) is the choice of the consistent estimator of the covariance matrix. $\mathbf{V}_{\mathbf{r}}^*(\hat{\boldsymbol{\eta}})$ contains the matrix $T^{-1}\sum\hat{\mathbf{v}}_t^*\hat{\mathbf{v}}_t^{*'}$ whereas its expectation $2\mathbf{I}_m$ appears in $\mathbf{V}_{\mathbf{r}}(\hat{\boldsymbol{\eta}})$. The two covariance matrices are thus asymptotically equal so that the statistics (3.5) and (3.6) and thus (3.3) and (3.5) have the same asymptotic null distribution. The reason for introducing (3.3) as Li and Mak (1994) already derived the same test is that obtaining the test using the LM principle makes it easy to see how the test can be made robust against nonnormal errors. Furthermore, the present test is can be viewed as a natural extension of Engle's (1982) classic LM test of no ARCH to this situation. It is well-known that the McLeod and Li (1983) portmanteau test and the Engle (1982) test are asymptotically equivalent, see, for example, Luukkonen et al. (1988b). When $h_t = \alpha_0$ in (2.2), the Li and Mak portmanteau test and our LM test collapse into the McLeod and Li (1983) and the Engle test, respectively.

Our test or the Li and Mak test are in principle general misspecification tests but, as is clear from above, they also have an LM interpretation, that is, they are LM tests against a certain parametric alternative. If the model builder testing his or her GARCH model believes that the only relevant alternative is a higher order GARCH model, then the LM test of Bollerslev (1986) is an obvious test to use. If a more general misspecification of the model is not excluded a priori, then the test of this section can be viewed as a useful complement to Bollerslev's test. Other types of misspecification will be considered next.

4. Misspecification of structure

In this section, we present two misspecification tests for an estimated conditional variance model. The first one is a test against nonlinearity or, in some cases, asymmetry. It is a modification of a test in Hagerud (1997). Second, we propose a test of parameter constancy against smooth continuous change in parameters. These tests may be viewed as conditional variance counterparts of the tests for the nonlinear conditional mean presented in Eitrheim and Teräsvirta (1996). To describe the common features in these tests we first introduce some notation and thereafter consider the two tests separately.

4.1. Notation

Consider now an augmented version of model (2.2),

$$\varepsilon_t = z_t(h_t + g_t)^{1/2}, \tag{4.1}$$

where $g_t = g(\boldsymbol{\pi}_1\boldsymbol{\pi}_2, \boldsymbol{\eta}; \tilde{\mathbf{s}}_t)$ such that $g(\mathbf{0}, \boldsymbol{\pi}_2, \boldsymbol{\eta}, \tilde{\mathbf{s}}_t) \equiv 0$, and $\tilde{\mathbf{s}}_t$ is a vector of random variables. The null hypothesis to be tested, using this notation, is $\boldsymbol{\pi}_1 = 0$. As an example, the alternative in the LM test of Bollerslev (1986) for testing a GARCH(p, q) model against GARCH$(p, q + r)$ or GARCH$(p + r, q)$, $r > 0$, alternatives, belongs to class (4.1). As another example, one could think of a "Ramsey-type" functional form LM test, in which case $\tilde{\mathbf{s}}_t = \mathbf{s}_t$, and $g_t = \pi_1(\boldsymbol{\eta}'\mathbf{s}_t)^2$ or $g_t = (\boldsymbol{\eta}'\mathbf{s}_t)[\exp\{\pi_1(\boldsymbol{\eta}'\mathbf{s}_t)\} - 1]$, where π_1 is a scalar and $\pi_2 = 0$. The null hypothesis would, as above, be $\pi_1 = 0$. This would lead to a one-degree-of-freedom test. When applying GARCH models the degrees of freedom are not usually a problem, however, so that we may consider more richly parameterized alternatives. We shall concentrate on tests of linearity and parameter constancy of the GARCH(p, q) model.

4.2. Testing linearity

When volatility in return series is modelled with GARCH models, we may sometimes expect the response to be a function not only of the size of the shock but also of the direction. Engle and Ng (1993), see also references therein, considered this possibility. We call such a response to a shock asymmetric and parameterize it by generalizing the GJR–GARCH model of Glosten et al. (1993). This is done in three ways. First, we make the transition between the extreme regimes smooth. Second, we incorporate a nonlinear version of the quadratic GARCH model of Sentana (1995) in our alternative. Finally, while the GJR–GARCH model is asymmetric, the present generalization may in some special cases remain symmetric, although the model is nonlinear. For smooth transition GARCH, see also Hagerud (1997), González-Rivera (1998) and Anderson et al. (1999). Our maintained model is in fact a special case of the model of Anderson et al. (1999). Let

$$H_n(x_t;\gamma,\mathbf{c}) = \left(1+\exp\left(-\gamma\prod_{l=1}^{n}(x_t-c_l)\right)\right)^{-1}, \quad \gamma>0,\ c_1 \leqslant \cdots \leqslant c_n, \tag{4.2}$$

where x_t is the transition variable at time t, γ is a slope parameter, and $\mathbf{c}=(c_1,\ldots,c_n)$ a location vector. Conditions $\gamma>0$ and $c_1 \leqslant \cdots \leqslant c_n$ are identifying conditions. When $\gamma=0$, $H_n(x_t;\gamma,\mathbf{c}) \equiv \frac{1}{2}$. Typically in practice, $n=1$ or $n=2$. The former choice yields a standard logistic function. When the slope parameter $\gamma \to \infty$, (4.2) with $n=1$ becomes a step function whose value equals one for $x_t \geqslant c_1$ and zero otherwise. When $n=2$, (4.2) is symmetric about $(c_1+c_2)/2$, its minimum value, achieved at this point, lies between zero and $\frac{1}{2}$, and the value of the function approaches unity as $x_t \to \pm\infty$. When $\gamma \to \infty$, function (4.2) becomes a "double step" function that obtains value zero for $c_1 \leqslant x_t < c_2$ and unity otherwise. The logistic function (4.2) is used for parameterizing the maintained model. Let $\bar{H}_n = H_n - \frac{1}{2}$. This transformation simplifies notation in deriving the test but does not affect the generality of the argument. The alternative to standard GARCH may now be written as (4.1) where

$$g_t = \sum_{j=1}^{q}\alpha_{0j}\bar{H}_n(\varepsilon_{t-j};\gamma,\mathbf{c}) + \sum_{j=1}^{q}\alpha_{1j}\bar{H}_n(\varepsilon_{t-j};\gamma,\mathbf{c})\varepsilon_{t-j}^2. \tag{4.3}$$

Using previous notation, $\boldsymbol{\pi}_1=\gamma$ and $\boldsymbol{\pi}_2=(c_1,\ldots,c_n,\alpha_{01},\ldots,\alpha_{0q},\alpha_{11},\ldots,\alpha_{1q})'$ in g_t, and $\tilde{\mathbf{s}}_t=\mathbf{s}_t$. The sufficient but not necessary conditions for $h_t+g_t>0$ are $\alpha_0>0, \sum_{i=1}^{q}\alpha_{0i}>0, \alpha_j>0, \alpha_j+\alpha_{1j}>0,\ j=1,\ldots,q$. Assuming $n=1, \alpha_{0j}=0,\ j=1,\ldots,q$, and letting $\gamma \to \infty$ in (4.3) yields the GJR–GARCH model. The test of the standard GARCH model against nonlinear GARCH in Hagerud (1997) may be viewed as a special case of this specification with $g_t=\sum_{j=1}^{q}\alpha_{1j}\bar{H}_n(\varepsilon_{t-j};\gamma,\mathbf{c})\varepsilon_{t-j}^2$.

Another way of parameterizing the alternative is to assume that the transition variable has a fixed delay. This assumption results in the following conditional variance model:

$$g_t = \alpha_{0d}\bar{H}_n(\varepsilon_{t-d};\gamma,\mathbf{c}) + \sum_{j=1}^{q}\alpha_{1j}\bar{H}_n(\varepsilon_{t-d};\gamma,\mathbf{c})\varepsilon_{t-j}^2. \tag{4.4}$$

This specification resembles the one in Teräsvirta (1994) for the STAR-type conditional mean. Probably the most common case in practice, however, is $d=q=p=1$ so that (4.3) and (4.4) coincide. Note that we do not impose any nonlinear structure on the h_{t-j}, $j=1,\dots,p$, as the alternative model is already very flexible without such an extension.

These smooth transition alternatives pose an identification problem. The null hypothesis can be expressed as $\mathrm{H}_0\colon \gamma=0$ in $\bar{H}_n$. It is seen that the remaining parameters in (4.2) or (4.3) are only identified under the alternative. In other words, the elements of $\boldsymbol{\pi}_2$ in g_t are nuisance parameters that do not appear in the null model. Thus, the classical test statistics are not available as their asymptotic null distributions are unknown; see, for example, Hansen (1996) for discussion. We circumvent the identification problem by following Luukkonen et al. (1988a). This is done by expanding the transition function $\bar{H}_n$ into a first-order Taylor series around $\gamma=0$, replacing the transition function with this Taylor approximation in (4.3) and rearranging terms. As a result,

$$h_t=(\boldsymbol{\eta}^*)'\mathbf{s}_t \quad \text{and} \quad g_t=\boldsymbol{\beta}_1'\mathbf{v}_{1t}+\sum_{i=3}^{n+2}\boldsymbol{\beta}_i'\mathbf{v}_{it}+R_1 \tag{4.5}$$

in (4.1) and, furthermore, $\boldsymbol{\beta}_i=(\beta_{i1},\dots,\beta_{iq})'=\gamma\tilde{\boldsymbol{\beta}}_i$, $\tilde{\boldsymbol{\beta}}_i\neq\mathbf{0}$, $\mathbf{v}_{it}=(\varepsilon_{t-1}^i,\dots,\varepsilon_{t-q}^i)'$, $i=1,3,\dots,n+2$, and R_1 is the remainder. The new null hypothesis, implied by (4.5), is $\mathrm{H}_0\colon \boldsymbol{\beta}_1=\boldsymbol{\beta}_3=\cdots=\boldsymbol{\beta}_{n+2}=\mathbf{0}$. Note that under $\mathrm{H}_0\colon R_1=0$, so that the stochastic remainder does not affect the distributional properties of the test statistic when the null hypothesis holds. With h_t and g_t defined as in (4.5), Eq. (4.1) may be called an "auxiliary GARCH equation". We make the following additional assumption:

(A.1) Under H_0, $\mathsf{E}\varepsilon_t^{2(n+2)}<\infty$.

Remark. The restrictions Assumption (A.1) implies on the parameters in the standard GARCH(1, 1) model appeared in Nelson (1990), see also He and Teräsvirta (1999). Recently, Carrasco and Chen (2002) worked out conditions for the existence of moments of higher order than four for the GARCH(p,q) model.

We define the conditional quasilog-likelihood for observation t as follows:

$$\ell_t=-(1/2)\ln(h_t+g_t)-\frac{\varepsilon_t^2}{2(h_t+g_t)}. \tag{4.6}$$

An LM test for the null hypothesis may now be derived starting from (4.6) along the same lines as in Appendix A. Note that g_t is not a function of $\boldsymbol{\eta}^*$ and that under H_0, $g_t\equiv 0$. Here, $\hat{\mathbf{x}}_t=\hat{h}_t^{-1}\partial\hat{h}_t/\partial\boldsymbol{\eta}^*$ ($\partial\hat{h}_t/\partial\boldsymbol{\eta}^*$ denotes $\partial(h_t+g_t)/\partial\boldsymbol{\eta}^*$ evaluated under H_0). In fact, $\hat{\mathbf{x}}_t$ has exactly the same form as in Theorem 3.1 because the null model is the GARCH(p,q) model in both cases. On the other hand, $\hat{\mathbf{v}}_t=(\hat{\mathbf{v}}_{1t}',\hat{\mathbf{v}}_{3t}',\dots,\hat{\mathbf{v}}_{n+2,t}')'$. Finally, Assumption (A.1) is needed for the existence of $\mathsf{E}\mathbf{v}_{n+2,t}\mathbf{v}_{n+2,t}'$.

Theorem 4.1. *Consider the auxiliary GARCH equation* (4.1) *where* $\{z_t\}$ *is a sequence of independent identically distributed random variables with zero mean, unit variance* *and* $\mathsf{E}z_t^3 = 0$, *with* h_t *and* g_t *are defined as in* (4.5). *Under the assumption* (A.1) *and* $\mathrm{H}_0\colon \boldsymbol{\beta}_1 = \boldsymbol{\beta}_3 = \cdots = \boldsymbol{\beta}_{n+2} = \mathbf{0}$ *in* (4.5), *the statistic* (3.3) *with* $\hat{\mathbf{x}}_t$ *as in Theorem* 3.1 *and* $\hat{\mathbf{v}}_t = (\hat{\mathbf{v}}_{1t}', \hat{\mathbf{v}}_{3t}', \ldots, \hat{\mathbf{v}}_{n+2,t}')'$ *with* $\mathbf{v}_{it} = (\varepsilon_{t-1}^i, \ldots, \varepsilon_{t-q}^i)', i = 1, 3, \ldots, n+2$, *has an asymptotic* χ^2 *distribution with* $m = (n+1)q$ *degrees of freedom.*

A similar result may be derived for the fixed delay case. It is also possible to test subhypotheses. For example, if the alternative to GARCH is the modified quadratic GARCH model in which only the intercept displays a smooth transition, the null hypothesis equals $\boldsymbol{\beta}_1 = \mathbf{0}$, whereas $\boldsymbol{\beta}_3 = \cdots = \boldsymbol{\beta}_{n+2} = \mathbf{0}$ even in the maintained model.

This test is a linearity test whose alternative generally implies asymmetry. Note, however, that if $n = 2$ and $c_1 = -c_2$ in (4.2), the smooth transition GARCH model is still symmetric. Such a model may be useful, for example, in situations where the large shocks die out so much more quickly than small shocks that the response cannot be adequately characterized by linear GARCH models.

The test may be carried out in the TR^2 form via an auxiliary regression exactly as in the previous section. The same is true for the robust version of the linearity test.

4.3. Testing parameter constancy

Testing parameter constancy is important in its own right but also because nonconstancy may manifest itself as an apparent lack of covariance stationarity (IGARCH); see, for example, Lamoureux and Lastrapes (1990). Here, we assume that the alternative to constant parameters in the conditional variance is that the parameters, or a subset of them, change smoothly over time. Lin and Teräsvirta (1994) applied this idea to testing parameter constancy in the conditional mean. Let the time-varying parameter $\boldsymbol{\eta}(t) = \boldsymbol{\eta} + \boldsymbol{\lambda}\bar{H}_n(t; \gamma, \mathbf{c})$. If the null hypothesis only concerns a subset of parameters then only the corresponding elements in $\boldsymbol{\lambda}$ are assumed to be nonzero a priori. Again, we define $\bar{H}_n = H_n - \frac{1}{2}$ where the transition function $H_n(t; \gamma, \mathbf{c})$ is assumed to be a logistic function of order n defined in (4.2) with $x_t \equiv t$. When $\gamma \to \infty$, $\bar{H}_1(t; \gamma, \mathbf{c})$ becomes a step-function and characterizes a single structural break in the model. Chu (1995) and Lin and Yang (1999) discussed testing parameter constancy against this alternative. The null hypothesis of parameter constancy becomes $\mathrm{H}_0\colon \gamma = 0$ under which $\boldsymbol{\eta}(t) \equiv \boldsymbol{\eta}$. By defining $\boldsymbol{\pi}_1 = \gamma$ and $\boldsymbol{\pi}_2 = (\boldsymbol{\lambda}', \mathbf{c}')'$ we can consider this as a special case of (4.1) with $g_t = (\boldsymbol{\lambda}'\mathbf{s}_t)\bar{H}_n(t; \gamma, \mathbf{c})$. We can again circumvent the lack of identification under the null hypothesis by a Taylor approximation of the transition function. A first-order Taylor-expansion of $\bar{H}_n(t; \gamma, \mathbf{c})$ around $\gamma = 0$ yields, after a reparameterization, model (4.1) with

$$h_t = (\boldsymbol{\eta}^*)'\mathbf{s}_t \quad \text{and} \quad g_t = \sum_{i=1}^{n} \boldsymbol{\beta}_i'\mathbf{v}_{it} + R_2. \tag{4.7}$$

In (4.7), $\boldsymbol{\beta}_i = \gamma\tilde{\boldsymbol{\beta}}_i, \tilde{\boldsymbol{\beta}}_i \neq \mathbf{0}, i = 1, \ldots, n$. The null hypothesis based on (4.7) equals $\boldsymbol{\beta}_1 = \cdots = \boldsymbol{\beta}_n = \mathbf{0}$, and the remainder $R_2 = 0$ when the null hypothesis is valid. Furthermore, $\mathbf{v}_{it} = t^i\mathbf{s}_t, i = 1, \ldots, n$. Define $\hat{\mathbf{v}}_{it} = t^i\hat{\mathbf{s}}_t, i = 1, \ldots, n$, and $\hat{\mathbf{v}}_t = (\hat{\mathbf{v}}_{1t}', \ldots, \hat{\mathbf{v}}_{nt}')'$.

The conditional quasilog-likelihood has the form (4.6), and we can now derive an LM test for parameter constancy using the definitions in (4.7). This yields the following result:

Theorem 4.2. *Consider the auxiliary GARCH equation* (4.1) *with* (4.7) *where* $\{z_t\}$ *is a sequence of independent identically distributed random variables with zero mean, unit variance and* $\mathsf{E}z_t^3 = 0$. *Under assumption* (A.2) *and* $\mathrm{H}_0\colon \boldsymbol{\beta}_1 = \cdots = \boldsymbol{\beta}_n = \mathbf{0}$ *in* (4.7), *the statistic* (3.3) *with* $\hat{\mathbf{x}}_t$ *as in Theorem* 3.1 *and* $\hat{\mathbf{v}}_t = (\hat{\mathbf{v}}_{1t}, \ldots, \hat{\mathbf{v}}_{nt})'$, $\hat{\mathbf{v}}_{it} = t^i \hat{\mathbf{s}}_t, i = 1, \ldots, n$, *has an asymptotic* χ^2 *distribution with* $m = n(p+q+1)$ *degrees-of-freedom.*

Remark 1. The proof requires the assumption $\mathsf{E}\mathbf{s}_t\mathbf{s}_t' < \infty$, which is equivalent to assuming $\mathsf{E}\varepsilon_t^4 < \infty$. Since the regularity conditions assumed throughout require even higher order moments, we do not explicitly mention this condition in the theorem.

Remark 2. A number of terms in the auxiliary GARCH equation now contain trending variables. Nevertheless, applying the results of Lai and Wei (1982) as when proving a corresponding result for linear conditional mean models, see Lin and Teräsvirta (1994), it can be shown that the asymptotic null distribution even in this case is a chi-squared one. This makes the test easy to apply.

Even this test may be carried out using an auxiliary regression as in Section 3.1. The procedure for carrying out the robust test, described in Section 3.1, is also valid for the parameter constancy test.

An advantage of a parametric alternative to parameter constancy is that if the null hypothesis is rejected we can estimate the parameters of the alternative model. This helps us find out where in the sample the parameters under test have been changing and how rapid the change appears to be. This is useful information when respecification of the model to achieve parameter constancy is attempted.

5. Simulation experiment

The above distribution theory is asymptotic, and we have to find out how our tests behave in finite samples. This is done by simulation. For all simulations we use the following data generating process (DGP)

$$y_t = \varepsilon_t,$$

where ε_t follows a standard GARCH process (size simulations) or one of the alternatives discussed above (power). The random numbers are generated by the random number generator in GAUSS 3.2.31. The first 200 observations of each generated series are always discarded to avoid initialization effects. All experiments are performed with series of 1000 observations. For each design, a total of 1000 replications have been carried out. The distribution for the error process $\{z_t\}$ is either standard normal or a standardized (unit variance) $t(d)$ distribution, where $d = 3, 4, 5$. We report results from experiments with normal and $t(3)$-errors. The ones with the other two t-distributions lie between these two extremes. Note that for $t(3)$ and $t(4)$-errors, the fourth moment of ε_t does not exist, which contradicts our assumptions. The idea is to investigate the effect of this violation of the conditions on the results.

Both the normality-based LM and the robust version of each test are considered. This is also true for the tests to which we compare our tests. A general result valid for practically all simulation experiments is that the nonrobust test is undersized when the error distribution is any one of the three t-distributions. Note that the results appearing in the figures are based on size-adjusted tests, so that this finding is not visible there. The robust tests are more powerful than the original LM or LM type tests. A general conclusion from our experiments is that in applied work one should apply the robust versions of our tests.

5.1. No remaining ARCH

First, we consider the test of no ARCH in the standardized errors. We define a DGP such that the conditional variance follows a GARCH(1,2) process:

$$h_t = 0.5 + 0.05\varepsilon_{t-1}^2 + \alpha\varepsilon_{t-2}^2 + 0.9h_{t-1}. \tag{5.1}$$

In the Monte Carlo experiment the values of α vary within limits such that the conditional variance of the process remains positive with probability 1, see Nelson and Cao (1992), and covariance stationarity holds. This is the case when $-0.045 < \alpha < 0.05$. For $\alpha = 0$ the DGP reduces to a standard GARCH(1,1) model. The results for the test of no ARCH in the standardized errors are reported in Fig. 1. The nominal significance level equals 0.1 (the results for levels 0.05 and 0.01 are not reported), and we use the test of Bollerslev (1986) against the GARCH(1,2) model as a benchmark. In the simulations, it was always assumed that the additional ARCH component in our test was ARCH(1), so that $\hat{\mathbf{v}}_t = \hat{\varepsilon}_{t-1}^2/\hat{h}_{t-1}$. Robust and nonrobust versions of both tests were simulated.

Our results indicate that the robust tests are well-sized, whereas the nonrobust ones are undersized when the error distribution is a t-distribution. The size-adjusted power results appearing in Fig. 1 are somewhat ambiguous in the normal case. The nonrobust tests are more powerful than the robust ones for negative values of α but less powerful for positive values of this parameter. This is true for both Bollerslev's tests and ours. When the errors are t-distributed, a clear difference in power in favour of the robust tests emerges. Note, however, that in that case both tests appear somewhat biased. The two tests in our comparison seem to have very similar size-adjusted power, although our test has somewhat greater power when the errors follow a $t(3)$-distribution. As Bollerslev's test is an LM test against the data-generating alternative, we may conclude that our test works very well in this experiment.

5.2. Testing linearity

Our linearity test can be expected to be powerful against smooth transition alternatives for which it is designed. Assessing its performance against, say, the GJR–GARCH

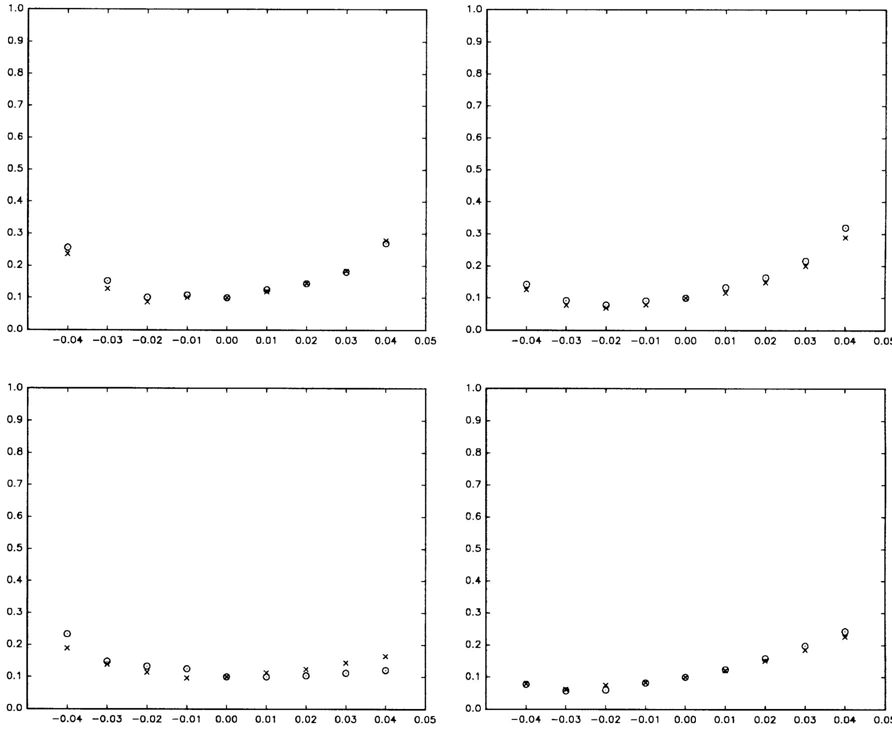

Fig. 1. Results from size and power simulations of the LM test of no remaining ARCH (circle) and the LM test of GARCH(1,1) against GARCH(1,2) (cross). DGP: GARCH(1,2), Eq. (5.1). x-axis: α, y-axis: size-adjusted power. Upper panel, left: normal errors, nonrobust test, upper panel, right: normal errors, robust test, lower panel, left: $t(3)$-errors, nonrobust test, lower panel, right: $t(3)$-errors, robust test. Number of observations=1000, number of replications=1000.

model would constitute a tougher trial for the test. This would also make it possible to compare the performance of the tests with the sign-bias and negative size-bias tests of Engle and Ng (1993). The DGP is, accordingly, (2.2) with

$$h_t = 0.005 + 0.23[|\varepsilon_{t-1}| + \omega\varepsilon_{t-1}]^2 + 0.7h_{t-1}, \tag{5.2}$$

where ε_t is assumed conditionally normal. In Eq. (5.2), $|\omega| < 0.551$ is required for covariance stationarity. This is the model Engle and Ng (1993) used in their simulation except that the coefficient of $[|\varepsilon_{t-1}| + \omega\varepsilon_{t-1}]^2$ equals 0.23 against 0.28 in Engle and Ng (1993). This change guarantees the existence of the unconditional fourth moment under the null hypothesis $\omega = 0$; the relevant moment condition appeared in He and Teräsvirta (1999). (The sign-bias and size-bias tests require a finite fourth moment.) The sign-bias and size-bias tests are designed for detecting asymmetry in the conditional variance. We compute the values of these test statistics using the quadratic form. The difference in results between this test and the TR^2 version, suggested by Engle and Ng (1993), is negligible at our sample size. Engle and Ng (1993) used (2.2) and (5.2)

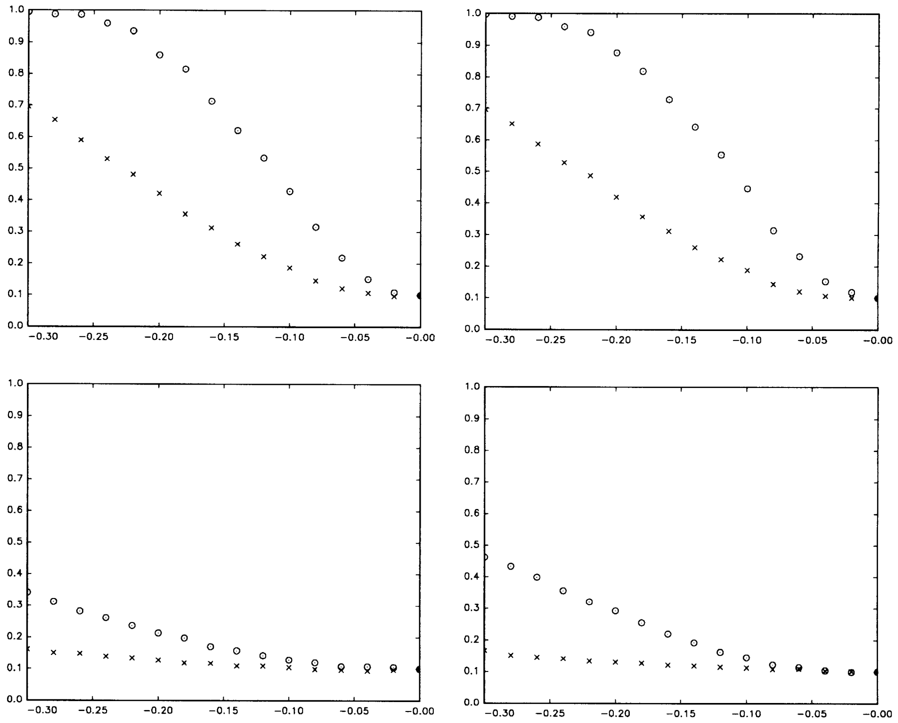

Fig. 2. Results from size and power simulations of the LM-type parameter constancy test (circle) and the robust version of the parameter constancy test of Chu (1995) (cross). DGP: GARCH(1,1) with a single structural break, equation (5.3). x-axis: η, y-axis: size-adjusted power. Upper panel, left: normal errors, nonrobust test, upper panel, right: normal errors, robust test, lower panel, left: $t(3)$-errors, nonrobust test, lower panel, right: $t(3)$-errors, robust test. Number of observations=1000, number of replications=1000.

with $\omega = -0.23$ as the DGP in their evaluation of the sign-bias test. We also considered robust versions of the two tests.

The results for the sign-bias and negative size-bias tests are rather similar. Thus, the latter test is excluded from Fig. 2 where our test is compared with the sign-bias test. The nonrobust sign-bias and our linearity test are undersized for the $t(3)$-errors. The robust versions of the tests have roughly the correct size.

As for size-adjusted power, results for $\omega < 0$ can be found in Fig. 2. For positive values of ω the results are similar and therefore omitted. Our smooth transition GARCH test was computed by assuming $n = 1$ in (4.2). When the errors are normal, 1000 observations are enough for the robust tests to perform as well as the nonrobust ones. In the case of t-distributed errors, the robust tests are always more powerful than the nonrobust ones. Our linearity test is clearly more powerful than the sign-bias test even if the alternative is a GJR–GARCH and not a genuine smooth transition GARCH model.

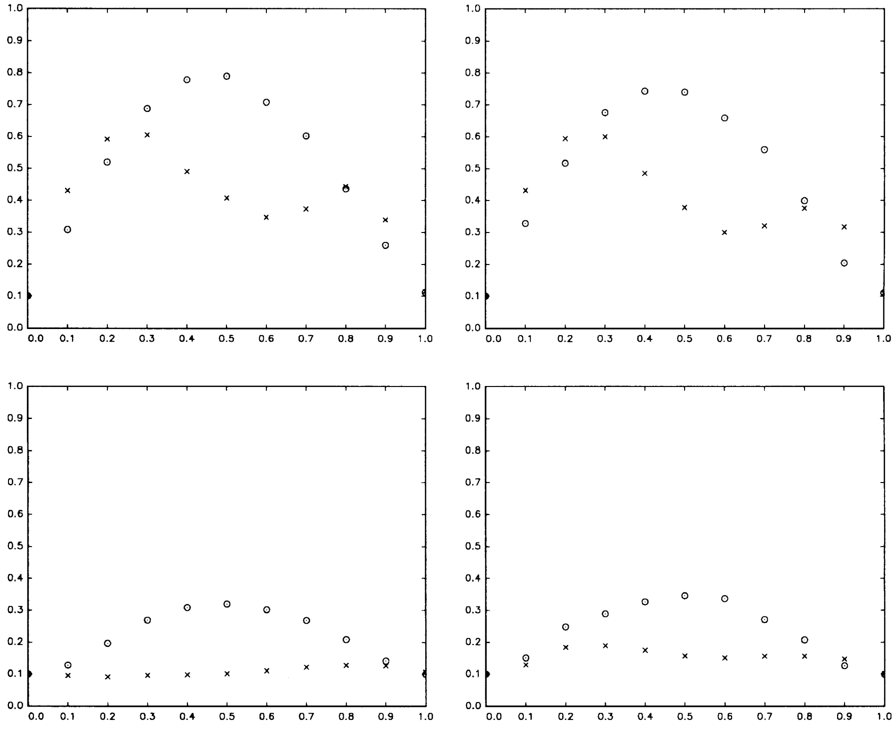

Fig. 3. Results from power simulations of the LM-type parameter constancy test (circle) and the parameter constancy test of Chu (1995) (cross). DGP: GARCH(1,1) with a double structural break, Eq. (5.3). x-axis: η, y-axis: size-adjusted power. Upper panel, left: normal errors, nonrobust test, upper panel, right: normal errors, robust test, lower panel, left: $t(3)$-errors, nonrobust test, lower panel, right: $t(3)$-errors, robust test. Number of observations=1000, number of replications=1000.

5.3. Testing parameter constancy

We consider two cases of parameter nonconstancy: the DGP is a GARCH(1,1) model with either (a) a single or (b) a double structural break. We did not simulate smooth parameter change because our test can be expected to perform well against such an alternative. Our choice of alternative gives us an opportunity to compare our test with tests against a single structural break, and we choose the test of Chu (1995) for the purpose. If T is the total number of observations, the single structural break parameterization (a) is assumed to have a break at time ηT where η lies between 0 to 1. The double structural break parameterization (b) postulates a change at time $\eta_1 T$ and a return to the original parameters at $\eta_2 T$, $0 \leqslant \eta_1 < \eta_2 \leqslant 1$. Our test is computed with $n = 1$, case (a), and $n = 2$, case (b), where n is the order of the logistic function in (4.2).

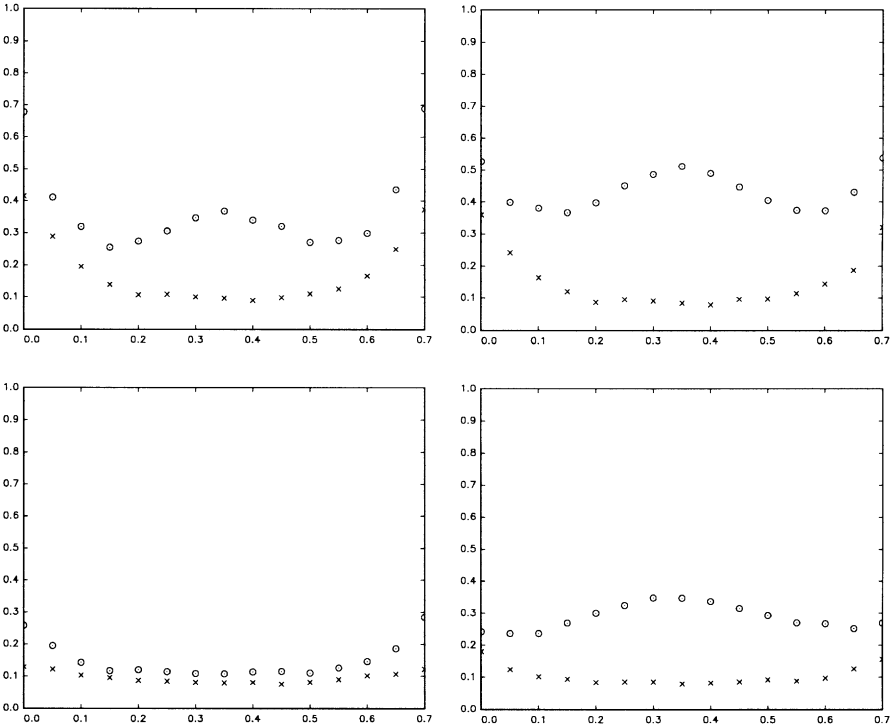

Fig. 4. Results from size and power simulations of the LM-type linearity test (circle) and the sign-bias test of Engle and Ng (1993) (cross). DGP: GJR–GARCH(1,1), Eq. (5.2). x-axis: ω, y-axis: size-adjusted power. Upper panel, left: normal errors, nonrobust test, upper panel, right: normal errors, robust test, lower panel, left: $t(3)$-errors, nonrobust test, lower panel, right: $t(3)$-errors, robust test. Number of observations=1000, number of replications=1000.

We consider the following model for a change in the constant term:

$$h_t = 0.5 + 0.1\varepsilon_{t-1}^2 + 0.8h_{t-1}, \quad \text{(a) } t < \eta T, \text{ (b) } t < \eta_1 T, t > \eta_2 T,$$

$$h_t = 0.5(1 + \Delta) + 0.1\varepsilon_{t-1}^2 + 0.8h_{t-1}, \quad \text{(a) } t \geqslant \eta T, \text{ (b) } \eta_1 T \leqslant t \leqslant \eta_2 T, \tag{5.3}$$

where $\Delta = 0.4, 0.8$. Chu (1995) used (a) in (5.3) as the DGP in his own simulation experiments. The power simulations for the DGPs in (5.3) with a single structural break at η for $\Delta = 0.8$ appear in Fig. 3. The values $\eta = 0, 1$ correspond to the null hypothesis. Even if only the intercept changes in (5.3), in simulating our LM-type test we assume that under the alternative, the break affects all three parameters. The asymptotic null distribution of our LM-type statistic is thus the $\chi^2(3)$-distribution.

We report results for $\Delta = 0.8$. The pattern of the results for $\Delta = 0.4$ is similar, only the power is lower. In another experiment, we allowed the coefficient of ε_{t-1}^2 to change once within the sample period. The behaviour of the tests was similar to the previous case, and no details are given here. In case (a) under normality, our LM-type test has the same as or higher power than the test of Chu (1995) for $0.3 \leqslant \eta \leqslant 0.7$, but otherwise the relationship is the opposite. When the errors follow a $t(3)$-distribution, the robust version of our test is more powerful than the standard version. It is also again more powerful than Chu's test for $0.3 \leqslant \eta \leqslant 0.7$. It seems, however, that Chu's robust test is oversized when the errors are $t(3)$-distributed whereas our robust one is not. On the other hand, in that situation the nonrobust LM-type test is clearly undersized (these two facts cannot be seen from Fig. 3).

We turn to case (b), the double structural break. The DGP in that design is such that $\eta_2 = \eta_1 + 0.3$ where η_1 varies from 0 to 0.7. Thus, for $\eta_1 = 0$ and $\eta_1 = 0.7$, the DGP only has a single structural break. Results of power simulations of this design with $\Delta = 0.8$ can be found in Fig. 4. In the case of a double break the test of Chu (1995) cannot be expected to be very powerful because the test is constructed for detecting just a single structural break. For this reason, its power is higher for η_1 close to zero and 0.7 than elsewhere. With few exceptions (nonrobust test, $t(3)$-distributed errors) our test with $n = 2$ does have size-adjusted power superior to that of Chu for all double break points considered. For t-distributed errors, the robust version of the test is clearly more powerful than the nonrobust one.

6. Conclusions

In this paper, we derive a number of misspecification tests for a standard GARCH (p, q) model. As all our test statistics are asymptotically χ^2-distributed under the null hypothesis, possible misspecification of a GARCH model can be detected at low computational cost. Because the tests of linearity and parameter constancy are parametric, the alternative may be estimated if the null hypothesis is rejected. This is useful for a model builder who wants to find out the possible weaknesses of the estimated specification. It may also give him/her useful ideas of how the model could be further improved.

We also highlight the fact that our test of no ARCH in the standardized error process is asymptotically equivalent to a portmanteau test of Li and Mak (1994). This links the work of these authors to our approach. As already mentioned, the advantage of our derivation of the no ARCH test statistic is that robustifying the test against nonnormal errors is straightforward.

Finally, the simulation results indicate that in practice, the robust versions of the tests should be preferred to nonrobust ones. At relevant sample sizes when the errors are normal, they are roughly as powerful as normality-based LM or LM type tests. When the errors are nonnormal, the robust tests are superior in terms of power to nonrobust ones. They can therefore be recommended as standard procedures when it comes to testing the adequacy of an estimated standard GARCH(p, q) model.

Acknowledgements

The first author acknowledges financial support from the Tore Browaldh's Foundation and Skandia Asset Management. The research of the second author has been supported by the Swedish Council for Research in Humanities and Social Sciences. We thank James Davidson and an anonymous referee for helpful suggestions and remarks but retain the responsibility for any errors and shortcomings in the paper.

Appendix A. Proof of Theorem 3.1

The conditional quasi log likelihood for observation t equals

$$\ell_t = -\frac{1}{2}\ln h_t - \frac{1}{2}\ln g_t - \frac{\varepsilon_t^2}{2h_t g_t}.$$

The partial derivative with respect to $\boldsymbol{\eta}$ is

$$\frac{\partial \ell_t}{\partial \boldsymbol{\eta}} = -\frac{1}{2}\left(\frac{1}{h_t}\frac{\partial h_t}{\partial \boldsymbol{\eta}} + \frac{1}{g_t}\frac{\partial g_t}{\partial \boldsymbol{\eta}} - \frac{\varepsilon_t^2}{h_t^2 g_t}\frac{1}{h_t}\frac{\partial h_t}{\partial \boldsymbol{\eta}} - \frac{\varepsilon_t^2}{h_t g_t^2}\frac{1}{g_t}\frac{\partial g_t}{\partial \boldsymbol{\eta}}\right)$$
$$= \frac{1}{2}(\varepsilon_t^2/h_t g_t - 1)\mathbf{x}_t + \frac{1}{2}(\varepsilon_t^2/h_t g_t - 1)\frac{1}{g_t}\frac{\partial g_t}{\partial \boldsymbol{\eta}},$$

where $\mathbf{x}_t = h_t^{-1}\partial h_t/\partial \boldsymbol{\eta}$ with $\partial h_t/\partial \boldsymbol{\eta} = \mathbf{s}_t + (\partial \mathbf{s}_t/\partial \boldsymbol{\eta}_t)'\boldsymbol{\eta}$, and $\partial g_t/\partial \boldsymbol{\eta} = (\partial \mathbf{v}_t/\partial \boldsymbol{\eta}_t)'\boldsymbol{\pi}$. Thus, $\partial g_t/\partial \boldsymbol{\eta}|_{\mathrm{H}_0} = 0$ so that $\partial \ell_t/\partial \boldsymbol{\eta}|_{\mathrm{H}_0} = (\frac{1}{2})(\varepsilon_t^2/h_t - 1)\mathbf{x}_t$. Likewise,

$$\frac{\partial \ell_t}{\partial \boldsymbol{\pi}} = \frac{1}{2}(\varepsilon_t^2/h_t g_t - 1)\frac{1}{g_t}\frac{\partial g_t}{\partial \boldsymbol{\pi}}$$

so that $\partial \ell_t/\partial \boldsymbol{\pi}|_{\mathrm{H}_0} = (\frac{1}{2})(\varepsilon_t^2/h_t - 1)\mathbf{v}_t$, where $\mathbf{v}_t = (\varepsilon_{t-1}^2/h_{t-1}, \ldots, \varepsilon_{t-m}^2/h_{t-m})'$. Let $\boldsymbol{\omega} = (\boldsymbol{\varphi}', \boldsymbol{\eta}')'$. The relevant component of the average pseudoscore evaluated at the true (under the null hypothesis) pair of parameter values $(\boldsymbol{\omega}_0, \mathbf{0})$ has the form

$$\mathbf{q}_T(\boldsymbol{\omega}_0, \mathbf{0}) = \frac{1}{2T}\sum_{t=1}^{T}(\varepsilon_t^2/h_t - 1)\begin{bmatrix}\mathbf{x}_t^0\\ \mathbf{v}_t\end{bmatrix},$$

where $\mathbf{x}_t^0$ equals $\mathbf{x}_t$ evaluated at $(\boldsymbol{\omega}_0, \mathbf{0})$. (We need not consider $(1/T)\sum_{t=1}^{T}\partial \ell_t/\partial \boldsymbol{\varphi}|_{\mathrm{H}_0}$, because assuming $\mathsf{E}z_t^3 = 0$ makes the covariance matrix of the full pseudoscore vector block diagonal.) Let $\hat{\varepsilon}_t = y_t - f(\mathbf{w}_t; \hat{\boldsymbol{\varphi}})$ where $\hat{\boldsymbol{\varphi}}$ is the quasi maximum likelihood estimator of $\boldsymbol{\varphi}$ under H_0. Furthermore, let $\hat{h}_t, \hat{\mathbf{x}}_t$ and $\hat{\mathbf{v}}_t$ equal $h_t, \mathbf{x}_t$ and $\mathbf{v}_t$ evaluated at $\boldsymbol{\omega} = \hat{\boldsymbol{\omega}}$. Under regularity conditions,

$$\sqrt{T}\mathbf{q}_T(\hat{\boldsymbol{\omega}}, \mathbf{0}) = \frac{1}{2\sqrt{T}}\sum_{t=1}^{T}(\hat{\varepsilon}_t^2/\hat{h}_t - 1)\begin{bmatrix}\hat{\mathbf{x}}_t\\ \hat{\mathbf{v}}_t\end{bmatrix}$$
$$= \begin{bmatrix}\mathbf{0}\\ (1/2\sqrt{T})\sum_{t=1}^{T}(\hat{\varepsilon}_t^2/\hat{h}_t - 1)\hat{\mathbf{v}}_t\end{bmatrix} \to \mathrm{N}(\mathbf{0}, \mathbf{J}_0),$$

in distribution as $T \to \infty$, where

$$\mathbf{J}_0 = \begin{bmatrix} \mathbf{J}_{011} & \mathbf{J}_{012} \\ \mathbf{J}_{021} & \mathbf{J}_{022} \end{bmatrix} = \mathsf{E}T^{-1}\mathbf{q}_T(\boldsymbol{\omega}_0, \mathbf{0})q_T(\boldsymbol{\omega}_0, \mathbf{0})'$$

$$= \frac{k}{4}\mathsf{E}\begin{bmatrix} \mathbf{x}_t^0\mathbf{x}_t^{0'} & \mathbf{x}_t^0\mathbf{v}_t' \\ \mathbf{x}_t^0\mathbf{v}_t' & \mathbf{v}_t\mathbf{v}_t' \end{bmatrix}.$$

with $k = \mathsf{E}(\varepsilon_t^2/h_t - 1)^2$. Thus,

$$LM_0 = \left(\sum_{t=1}^{T}(\hat{\varepsilon}_t^2/\hat{h}_t - 1)\hat{\mathbf{v}}_t')\mathbf{J}_{022}^{-1}\left(\sum_{t=1}^{T}(\hat{\varepsilon}_t^2/\hat{h}_t - 1\right)\hat{\mathbf{v}}_t'\right) \tag{A.1}$$

has an asymptotic χ^2 distribution with m degrees of freedom. Replacing $\mathbf{J}_{022}^{-1}$ in (A.1) by a consistent estimator

$$V(\hat{\boldsymbol{\eta}})^{-1} = (4T/\hat{k})\left\{\sum_{t=1}^{T}\hat{\mathbf{v}}_t\hat{\mathbf{v}}_t' - \sum_{t=1}^{T}\hat{\mathbf{v}}_t\hat{\mathbf{x}}_t'\left(\sum_{t=1}^{T}\hat{\mathbf{x}}_t\hat{\mathbf{x}}_t'\right)^{-1}\sum_{t=1}^{T}\hat{\mathbf{x}}_t\hat{\mathbf{v}}_t'\right\}^{-1}$$

with $\hat{k} = (1/T)\sum_{t=1}^{T}(\hat{\varepsilon}_t^2/\hat{h}_t - 1)^2$, yields the result; see Davidson (2000, Section 12.3.3). □

References

Anderson, H.M., Nam, K., Vahid, F., 1999. Asymmetric nonlinear smooth transition GARCH models. In: Rothman, P. (Ed.), Nonlinear time series analysis of economic and financial data. Kluwer, Boston.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327.

Carrasco, M., Chen, X., 2002. Mixing and moment properties of various GARCH and stochastic volatility models. Econometric Theory 18, 17–39.

Chu, C.-S.J., 1995. Detecting parameter shift in GARCH models. Econometric Reviews 14, 241–266.

Comte, F., Lieberman, O., 2000. Asymptotic theory for multivariate GARCH processes. Journal of Multivariate Analysis, forthcoming.

Davidson, J., 2000. Economic Theory. Blackwell, Oxford.

Demos, A., Sentana, E., 1998. Testing for GARCH effects: a one-sided approach Journal of Econometrics 86, 97–127.

Eitrheim, Ø., Teräsvirta, T., 1996. Testing the adequacy of smooth transition autoregressive models. Journal of Econometrics 74, 59–75.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 987–1007.

Engle, R.F., Ng, V.K., 1993. Measuring and testing the impact of news on volatility. Journal of Finance 48, 1749–1777.

Glosten, L., Jagannathan, R., Runkle, D., 1993. On the relation between expected value and the volatility of the nominal excess return on stocks. Journal of Finance 48, 1779–1801.

González-Rivera, G., 1998. Smooth transition GARCH models. Studies in Nonlinear Dynamics and Econometrics 3, 61–78.

Hagerud, G.E., 1997. A New Non-Linear GARCH Model. EFI, Stockholm School of Economics.

Hansen, B.E., 1996. Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica 64, 413–430.

He, C., Teräsvirta, T., 1999. Properties of moments of a family of GARCH processes. Journal of Econometrics 92, 173–192.

Lai, T.Z., Wei, C.Z., 1982. Least squares estimates in stochastic regression models with applications in identification and control of dynamic systems. Annals of Statistics 10, 154–166.

Lamoureux, C.G., Lastrapes, W.G., 1990. Persistence in variance, structural change and the GARCH model. Journal of Business and Economic Statistics 8, 225–234.

Lee, J., King, M., 1993. A locally most mean powerful based score test for ARCH and GARCH regression disturbances. Journal of Business and Economic Statistics 11, 17–27.

Li, W.K., Mak, T.K., 1994. On the squared residual autocorrelations in non-linear time series with conditional heteroskedasticity. Journal of Time Series Analysis 15, 627–636.

Lin, C.-F., Teräsvirta, T., 1994. Testing the constancy of regression parameters against continuous structural change. Journal of Econometrics 62, 211–228.

Lin, S.-J., Yang, J., 1999. Testing shift in financial models with conditional heteroskedasticity: an empirical distribution function approach. Research Paper 30, University of Technology, Quantitative Finance Research Group, Sydney.

Luukkonen, R., Saikkonen, P., Teräsvirta, T., 1988a. Testing linearity against smooth transition autoregressive models. Biometrika 75, 491–499.

Luukkonen, R., Saikkonen, P., Teräsvirta, T., 1988b. Testing linearity in univariate time series models. Scandinavian Journal of Statistics 15, 161–175.

McLeod, A.I., Li, W.K., 1983. Diagnostic checking ARMA time series models using squared residual autocorrelations. Journal of Time Series Analysis 4, 269–273.

Nelson, D.B., 1990. Stationarity and persistence in the GARCH(1,1) model. Econometric Theory 6, 318–334.

Nelson, D.B., Cao, C.Q., 1992. Inequality constraints in the univariate GARCH model. Journal of Business and Economic Statistics 10, 229–235.

Sentana, E., 1995. Quadratic ARCH models. Review of Economic Studies 62, 639–661.

Teräsvirta, T., 1994. Specification, estimation, and evaluation of smooth transition autoregressive models. Journal of the American Statistical Association 89, 208–218.

Wooldridge, J., 1991. On the application of robust, regression-based diagnostics to models of conditional means and conditional variances. Journal of Econometrics 47, 5–46.